Conference Call 4/26/01

Access to NPIRS DB

DB Tools: Karen and Stephanie have access to the production database and have DB2 Tools installed on their systems. Stephanie will initiate a NPIRS task to get Debra, Linda Q, Roger, and Patricia access to both databases and installation of the Tools.

SAS/DB2 Connect: Karen has gotten SAS Vn. 8.2 successfully installed with full connectivity and will assist in getting the software installed for HQ and Patricia. The plan is to use testdb (last updated 4/17) to run process logic testing and access the production database only when we're ready to run a report requiring current data. <u>Task</u>: Karen will document installation and configuration of the SAS software and the process for obtaining DB2 Tools. Reminder to the team: before anyone runs a process against the production database they are to check with Stephanie or Karen for its potential impact on the NPIRS processing load.

Regular reporting on changes to production database: Ask Sue to decide how and when she believes it best to do this report.

Access policy: Debra has obtained models of such policies and will document recommendations to submit to Edna for a decision. Suggested strategy is to determine first an initial policy for providing access to a smaller group of stat officers and national/headquarters Epi program people that would include handling cross-area access and identifying what they can do with the data (e.g., how adhere to confidentiality by areas/sites, statistical anonymity). The second phase would address progressively more authorized users until all who require access have it. Development of the policy is a headquarters task; the Team will handle its technical implementation.

Documentation

User Pops process document being prepared. Once the draft is complete, will distribute to rest of group (Debra, Patricia) for review. Workload processing logic is next to be prepared. It will include how data is handled at the PCC repository (what does/doesn't get exported), then at NPIRS (what processing takes place and the criteria by which a record gets counted for workload).

<u>Step 1</u>: Clarify and document the NPIRS logic for determining workload reportable (include unduplication process/algorithm).

<u>Step 2</u>: We will ask Melvin to document how the raw data received from the local PCC export process is handled at NPIRS, including how it is processed and eventually placed in the various NPIRS files.

<u>Step 3</u>: We will ask Lori to begin to document the programming logic she uses within the Statistical Record Export Program to determine which visits to send to NPIRS

<u>Step 4</u>: We will then take all of this logic to Lori Butcher (PCC) requesting that she make the local PCC Management Reports versions of the respective reports consistent with all of these (may require a iteratively versioning process).

Timing suggestion: at each user pop iteration, we note changes to logic and every 6-12 months take these changes to Lori for another patch to the appropriate PCC Management Reports report logics.

User Pops

CDMIS data: CDMIS data has been excluded from User Pops because not every facility has the data to export and therefore couldn't participate equally. Rus has received a request from the Urban programs that this decision be reconsidered because this disproportionately affects them. This is an issue for the program side of the house to decide and so it will be passed to Linda and Edna for a decision. <u>Task</u>: Stan to email Linda and Edna.

Regarding suspected gaps in data between '97-'99: User pops do not need to depend on the workload being verified. According to Patricia, this verification process is relatively new and was implemented in response to data quality questions. More important than basing user pops on verified workload is the iterative running of user pops for review and identification of issues, enabling us to put improved processes in place. We need to produce workloads and user pops routinely for review and secondarily decide when to pick a user pop for official approval. This running user pops iteratively is for us to review and identify issues and to enable us to put processes in place.

Regarding the Patient Reg Re-export: We will continue to work toward this but will not hold up production of our first internal user pop/workload until this is completed. Using Automatch or some other unduplication approach, we should be able to come up with something that has integrity. With agreement on this among team members, Debra will discuss it with Linda Querec because the program side should make this call.

Concurrency of User Pop/Workload Verification: We will work workload verification and user pop concurrently for some time. The decision to fill in holes of missing data will be made as we go along based on estimated impact – Stan will work on estimating the extent and impact of missing data (97-99). Patricia will work on buy-in to this iterative and concurrent approach to workload and user pops.

Unduplication: Karen has the files ready to run Automatch. We are trying to account for each user (once) by Area because the allocation process is by Area and Areas allocate internally to sites. We don't now have a certified method to 'undupe'. We need a certified process to use regularly for unduplication. Debra will check with Linda and Edna to make sure we have the authority to determine the unduplication criteria for regular use. Karen and Patricia will come up with a method and present it to the Team for approval.

Other Topics

Processing Data: Unprocessed patient reg data processed on 4/17. Need to prepare to schedule regular processing in NPIRS to start up after the first iteration of user pops.

Backup Procedure and Schedule: This has been implemented with the production database with offsite storage.

SAS contract: Joan will check on this status; anticipate their involvement to begin later in May or early June. We need to know what documentation needs to be in place for them.